

WEST Search History

[Hide Items](#)[Restore](#)[Clear](#)[Cancel](#)

DATE: Tuesday, September 20, 2005

Hide?	Set Name	Query	Hit Count
		<i>DB=PGPB,USPT,EPAB; PLUR=YES; OP=ADJ</i>	
<input type="checkbox"/>	L13	L3 and L9	0
<input type="checkbox"/>	L12	L9 and L11	0
<input type="checkbox"/>	L11	L7.ab.	410
<input type="checkbox"/>	L10	L9 and L7	29
<input type="checkbox"/>	L9	(536/23.1)![CCLS]	12263
<input type="checkbox"/>	L8	L7 and L1	1
<input type="checkbox"/>	L7	aldose reducta\$	1812
<input type="checkbox"/>	L6	aldose reducta?	1
<input type="checkbox"/>	L5	TTGTGGCCCACTTCCAGATCGAGAAGCTCT	0
<input type="checkbox"/>	L4	L3 and L1	1
<input type="checkbox"/>	L3	L2.ab.	52
<input type="checkbox"/>	L2	ARL	1208
<input type="checkbox"/>	L1	dai.in.	2900

END OF SEARCH HISTORY

us-10-653-681b-1.Oligo.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 12:07:39 ; Search time 698 Seconds
(without alignments)
9244.295 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaataataatcat 1090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	780	71.6	1337	5	AAS68608	Aas68608 DNA encod
2	780	71.6	1337	10	ADD71032	Add71032 Human ald
3	773	70.9	1549	12	ADK70274	Adk70274 Respirato
4	729	66.9	1508	3	AAC98140	Aac98140 Human col
5	729	66.9	1560	12	ADJ75119	Adj75119 Marker ge
6	729	66.9	1560	12	ADN04246	Adn04246 Antipsori
7	729	66.9	1560	13	ACN38728	Acn38728 Tumour-as
8	729	66.9	1560	13	ADS85007	Ads85007 Human ato
9	664	60.9	1316	5	AAF68405	Aaf68405 Human lun
10	664	60.9	1316	6	ABK38316	Abk38316 cDNA enco

us-10-653-681b-1.01igo.rng

	11	664	60.9	1316	7	ADS73134	Ads73134	Human kid
	12	664	60.9	1316	8	ACA10645	Aca10645	Human lun
	13	664	60.9	1316	8	ABX99596	Abx99596	Lung canc
	14	664	60.9	1316	10	ADH45842	Adh45842	Human lun
	15	664	60.9	1316	12	ADE72379	Ade72379	Human lun
	16	664	60.9	1316	13	ADJ19761	Adj19761	Human lun
	17	529	48.5	770	13	ADR98739	Adr98739	Lung spec
	18	529	48.5	1621	12	ADH13722	Adh13722	Human ENZ
c	19	389	35.7	558	10	ABZ84625	Abz84625	Toxicolog
	20	328	30.1	364	4	AAS39335	Aas39335	Novel hum
	21	316	29.0	585	2	AAZ24592	Aaz24592	Human lun
	22	316	29.0	585	3	AAC65831	Aac65831	Human lun
	23	316	29.0	585	6	ABL49050	Ab149050	Human lun
	24	316	29.0	585	6	ABQ92236	Abq92236	Human lun
	25	316	29.0	585	9	ADA28651	Ada28651	Human lun
	26	316	29.0	585	10	ADE53611	Ade53611	Human lun
	27	316	29.0	585	10	ADH36746	Adh36746	Human lun
	28	316	29.0	585	12	ADM56549	Adm56549	Human lun
	29	316	29.0	585	12	ADN89593	Adn89593	Human lun
c	30	316	29.0	857	9	ADA28650	Ada28650	Human lun
c	31	316	29.0	858	2	AAZ24591	Aaz24591	Human lun
c	32	316	29.0	858	3	AAC65830	Aac65830	Human lun
c	33	316	29.0	858	6	ABL49049	Ab149049	Human lun
c	34	316	29.0	858	6	ABQ92235	Abq92235	Human lun
c	35	316	29.0	858	10	ADE53610	Ade53610	Human lun
c	36	316	29.0	858	10	ADH36745	Adh36745	Human lun
c	37	316	29.0	858	12	ADM56548	Adm56548	Human lun
c	38	316	29.0	858	12	ADN89592	Adn89592	Human lun
	39	304	27.9	356	4	AAS39333	Aas39333	Novel hum
	40	252	23.1	1396	5	AAS91091	Aas91091	DNA encod
	41	249	22.8	861	5	AAS68606	Aas68606	DNA encod
	42	235	21.6	830	4	AAI92428	Aai92428	Human pol
	43	183	16.8	540	12	ADP28822	Adp28822	Human sec
	44	174	16.0	857	13	ADR98738	Adr98738	Lung spec
c	45	159	14.6	198	12	ACH84956	Ach84956	Human gen

RESULT 9

AAF68405

ID AAF68405 standard; cDNA; 1316 BP.

XX

AC AAF68405;

XX

DT 12-APR-2001 (first entry)

XX

DE Human lung tumour protein related nucleotide sequence SEQ ID NO:323.

XX

KW Human; lung cancer; lung tumour; lung tumour protein; gene therapy;
 KW lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
 KW cytostatic; antisense inhibition; ss.

XX

OS Homo sapiens.

XX

PN WO200100828-A2.

XX

PD 04-JAN-2001.

XX

PF 30-JUN-2000; 2000WO-US018061.

XX

PR 30-JUN-1999; 99US-00346492.

PR 15-OCT-1999; 99US-00419356.

PR 17-DEC-1999; 99US-00466867.

PR 30-DEC-1999; 99US-00476300.

PR 06-MAR-2000; 2000US-00519642.
 PR 22-MAR-2000; 2000US-00533077.
 PR 10-APR-2000; 2000US-00546259.
 PR 27-APR-2000; 2000US-00560406.
 PR 05-JUN-2000; 2000US-00589184.

XX

PA

(CORI-) CORIXA CORP.

XX

PI

Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
 Retter MW, Mannion J;

XX

DR

WPI; 2001-071488/08.

XX

PT

Lung tumor-associated proteins and the nucleic acids that encode them,
 useful for preventing, diagnosing and treating lung cancer.

XX

PS

Example 1; Page 249-250; 436pp; English.

XX

CC

The present invention describes immunogenic portions of lung tumour-associated proteins (I) and the nucleic acids (NAs) that encode them. (I) have cytostatic activity and can be used in gene therapy, antisense inhibition and in vaccines. The NAs and the lung tumour-associated proteins they encode may be used in the prevention, treatment and diagnosis of diseases associated with their inappropriate expression, especially lung cancers. For example, the NAs may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of the protein by expressing inactive proteins or to supplement the patients own production of (I). Additionally, the NAs may be used to produce the lung-tumour associated protein, according to standard recombinant DNA methodology. Conversely, antisense NA molecules may be administered to down regulate protein expression by binding with the cells own genes and preventing their expression. The NA and complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar NA sequences in samples, and hence which patients may be in need of treatment for lung cancer. The (I) may be used as antigens in the production of antibodies and in assays to identify modulators (agonists and antagonists) of the expression and activity of the protein. AAF68083 to AAF68878 and AAB76848 to AAB76878 represent human lung tumour protein related nucleotide and protein sequences which are used in the exemplification of the present invention

XX

SQ

Sequence 1316 BP; 385 A; 299 C; 308 G; 324 T; 0 U; 0 Other;

Query Match 60.9%; Score 664; DB 5; Length 1316;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 764; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 311 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACAGTGACTAA 370

Db 528 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACAGTGACTAA 587

Qy 371 CCAGGTTGAGTGTCACCCATACCTCACGCAGGAGAACTGATCCAGTACTGCCACTCCAA 430

Db 588 CCAGGTTGAGTGTCACCCATACCTCACACAGGAGAACTGATCCAGTACTGCCACTCCAA 647

Qy 431 GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC 490

Db 648 GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC 707

Qy 491 AGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA 550

Db 708 AGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA 767

us-10-653-681b-1.Oligo.rng

```

QY      551 AACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA 610
      |||
Db      768 AACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA 827
      |||
QY      611 GTCTGTGACACCAGCACGCATTGTTGAGAACATTCAGGTCTTTGACTTTAAATTGAGTGA 670
      |||
Db      828 GTCTGTGACACCAGCACGCATTGTTGAGAACATTCAGGTCTTTGACTTTAAATTGAGTGA 887
      |||
QY      671 TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA 730
      |||
Db      888 TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA 947
      |||
QY      731 ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG 790
      |||
Db      948 ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG 1007
      |||
QY      791 AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT 850
      |||
Db      1008 AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT 1067
      |||
QY      851 TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACGAGAATCGAGG 910
      |||
Db      1068 TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACGAGAATCGAGG 1127
      |||
QY      911 TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA 970
      |||
Db      1128 TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA 1187
      |||
QY      971 TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA 1030
      |||
Db      1188 TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA 1247
      |||
QY      1031 AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAA 1076
      |||
Db      1248 AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAA 1293
      |||

```

RESULT 10

ABK38316

ID ABK38316 standard; cDNA; 1316 BP.

XX

AC ABK38316;

XX

DT 21-MAY-2002 (first entry)

XX

DE cDNA encoding clone #18973 (L516S) of lung tumour protein.

XX

KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;
gene; ss.

XX

OS Homo sapiens.

XX

PN WO200204514-A2.

XX

PD 17-JAN-2002.

XX

PF 10-JUL-2001; 2001WO-US022058.

XX

PR 11-JUL-2000; 2000US-00614124.

PR 29-AUG-2000; 2000US-00651563.

PR 08-SEP-2000; 2000US-00658824.

PR 26-SEP-2000; 2000US-00671325.

PR 06-OCT-2000; 2000US-00677419.
 PR 30-OCT-2000; 2000US-00702705.
 PR 13-DEC-2000; 2000US-00736457.
 PR 03-MAY-2001; 2001US-00849626.

XX

PA (CORI-) CORIXA CORP.

XX

PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
 PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
 PI Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;

XX

DR WPI; 2002-164634/21.

XX

PT Novel polynucleotide encoding a lung tumor polypeptide useful for
 PT stimulating and/or expanding T cells specific for a tumor protein.

XX

PS Example 1; SEQ ID NO 323; 223pp; English.

XX

CC The invention describes an isolated polynucleotide and polypeptide useful
 CC for stimulating and/or expanding T cells specific for a tumour protein
 CC for determining the presence of a cancer in a patient. A composition
 CC containing the polynucleotide and/or polypeptide is useful for treating a
 CC lung cancer in a patient. The polypeptide is useful for removing tumour
 CC cells from a biological sample. The polynucleotide is also useful as
 CC probe or primer to detect the level of mRNA encoding a tumour protein.
 CC This sequence encodes a lung tumour associated protein or protein
 CC fragment, described in the method of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1316 BP; 385 A; 299 C; 308 G; 324 T; 0 U; 0 Other;

Query Match 60.9%; Score 664; DB 6; Length 1316;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 764; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 311 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACCACTGACTAA 370
 |||
 Db 528 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACCACTGACTAA 587
 Qy 371 CCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATCCAGTACTGCCACTCCAA 430
 |||
 Db 588 CCAGGTTGAGTGTACCCATACCTCACACAGGAGAACTGATCCAGTACTGCCACTCCAA 647
 Qy 431 GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC 490
 |||
 Db 648 GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC 707
 Qy 491 AGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA 550
 |||
 Db 708 AGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA 767
 Qy 551 AACC GCAGCC CAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA 610
 |||
 Db 768 AACC GCAGCC CAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA 827
 Qy 611 GTCTGTGACACCAGCACGCATTGTTGAGAACATT CAGGTCTTTGACTTTAAATTGAGTGA 670
 |||
 Db 828 GTCTGTGACACCAGCACGCATTGTTGAGAACATT CAGGTCTTTGACTTTAAATTGAGTGA 887
 Qy 671 TGAGGAGATGGCAACCATACTCAGCTTCAACAGAAACTGGAGGGCCTGTAACGTGTTGCA 730
 |||

us-10-653-681b-1.Oligo.rng

Db	888	TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA	947
Qy	731	ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG	790
Db	948	ATCCTCTCATTTGGAAGACTATCCCTTCAATGCAGAATATTGAGGTTGAATCTCCTGGTG	1007
Qy	791	AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	850
Db	1008	AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	1067
Qy	851	TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACGAGAATCGAGG	910
Db	1068	TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACGAGAATCGAGG	1127
Qy	911	TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA	970
Db	1128	TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA	1187
Qy	971	TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	1030
Db	1188	TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	1247
Qy	1031	AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAA	1076
Db	1248	AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAA	1293

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 11:49:49 ; Search time 5147 Seconds
(without alignments)
10261.548 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	780	71.6	1336	6	CQ718316	CQ718316 Sequence
2	780	71.6	1337	9	HSU37100	U37100 Homo sapien

3	729	66.9	1551	9	BC008837	BC008837	Homo sapi
4	729	66.9	1560	6	CQ776685	CQ776685	Sequence
5	721	66.1	1611	9	AF524864	AF524864	Homo sapi
6	664	60.9	1316	6	AR272611	AR272611	Sequence
7	664	60.9	1316	6	AR276192	AR276192	Sequence
8	664	60.9	1316	6	AR406467	AR406467	Sequence
9	664	60.9	1316	6	AR440317	AR440317	Sequence
10	664	60.9	1316	6	AR472475	AR472475	Sequence
11	664	60.9	1316	6	AR543128	AR543128	Sequence
12	664	60.9	1316	6	AX062696	AX062696	Sequence
13	664	60.9	1316	6	AX367613	AX367613	Sequence
14	664	60.9	1316	9	AF052577	AF052577	Homo sapi
15	574	52.7	574	9	AF044961	AF044961	Homo sapi
16	410	37.6	951	9	BT006794	BT006794	Homo sapi
17	410	37.6	951	12	BT007750	BT007750	Synthetic
18	409	37.5	948	9	CR541801	CR541801	Homo sapi
19	331	30.4	163631	9	AC009276	AC009276	Homo sapi
c 20	331	30.4	170919	9	AC078847	AC078847	Homo sapi
c 21	331	30.4	177373	2	AP002452	AP002452	Homo sapi
22	331	30.4	196039	2	AC055757	AC055757	Homo sapi
23	328	30.1	364	6	AX247463	AX247463	Sequence
24	316	29.0	585	6	AR176414	AR176414	Sequence
25	316	29.0	585	6	BD226027	BD226027	Compound
26	316	29.0	585	6	BD275698	BD275698	COMPOUNDS
27	316	29.0	585	6	AR220483	AR220483	Sequence
28	316	29.0	585	6	AR255477	AR255477	Sequence
29	316	29.0	585	6	AR281047	AR281047	Sequence
30	316	29.0	585	6	AR437838	AR437838	Sequence
31	316	29.0	585	6	AR476374	AR476374	Sequence
32	316	29.0	585	6	AR486565	AR486565	Sequence
33	316	29.0	585	6	AR541068	AR541068	Sequence
34	316	29.0	585	6	AX365699	AX365699	Sequence
c 35	316	29.0	858	6	AR176413	AR176413	Sequence
c 36	316	29.0	858	6	BD226026	BD226026	Compound
c 37	316	29.0	858	6	BD275697	BD275697	COMPOUNDS
c 38	316	29.0	858	6	AR220482	AR220482	Sequence
c 39	316	29.0	858	6	AR255476	AR255476	Sequence
c 40	316	29.0	858	6	AR281046	AR281046	Sequence
c 41	316	29.0	858	6	AR437837	AR437837	Sequence
c 42	316	29.0	858	6	AR476373	AR476373	Sequence
c 43	316	29.0	858	6	AR486564	AR486564	Sequence
c 44	316	29.0	858	6	AR541067	AR541067	Sequence
c 45	316	29.0	858	6	AX365698	AX365698	Sequence

ALIGNMENTS

RESULT 1
 CQ718316
 LOCUS CQ718316 1336 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 4250 from Patent WO02068579.
 ACCESSION CQ718316
 VERSION CQ718316.1 GI:42279173
 KEYWORDS .

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
 TITLE Kits, such as nucleic acid arrays, comprising a majority of
 humanexons or transcripts, for detecting expression and other uses
 thereof
 JOURNAL Patent: WO 02068579-A 4250 06-SEP-2002;
 PE Corporation (NY) (US)
 FEATURES Location/Qualifiers
 source 1. .1336
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 71.6%; Score 780; DB 6; Length 1336;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps
 0;

Qy 311 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACAGTGACTAA 370
 |||
 Db 557 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACAGTGACTAA 616
 Qy 371 CCAGGTTGAGTGTCAACCATACCTCACGCAGGAGAAACTGATCCAGTACTGCCACTCCAA 430
 |||
 Db 617 CCAGGTTGAGTGTCAACCATACCTCACGCAGGAGAAACTGATCCAGTACTGCCACTCCAA 676
 Qy 431 GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC 490
 |||
 Db 677 GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC 736
 Qy 491 AGAAGACCCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA 550
 |||
 Db 737 AGAAGACCCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA 796
 Qy 551 AACCGCAGCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA 610
 |||
 Db 797 AACCGCAGCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA 856
 Qy 611 GTCTGTGACACCAGCACGCATTGTTGAGAACATTTCAGGTCTTTGACTTTAAATTGAGTGA 670
 |||
 Db 857 GTCTGTGACACCAGCACGCATTGTTGAGAACATTTCAGGTCTTTGACTTTAAATTGAGTGA 916
 Qy 671 TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA 730
 |||
 Db 917 TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA 976
 Qy 731 ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG 790
 |||
 Db 977 ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG
 1036

Qy	791	AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	850
Db	1037	AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	
1096			
Qy	851	TTTAGCCAAGCTTATTTAAGATCACAGTGAAGTTAGTCCTGTTATAGACGAGAATCGAGG	910
Db	1097	TTTAGCCAAGCTTATTTAAGATCACAGTGAAGTTAGTCCTGTTATAGACGAGAATCGAGG	
1156			
Qy	911	TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAAATATCACAGAAAAGCA	970
Db	1157	TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAAATATCACAGAAAAGCA	
1216			
Qy	971	TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	
1030			
Db	1217	TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	
1276			
Qy	1031	AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAAATAATAATCAT	
1090			
Db	1277	AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAAATAATAATCAT	
1336			

us-10-653-681b-1.Oligo.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 12:07:39 ; Search time 698 Seconds
(without alignments)
9244.295 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	780	71.6	1337	5 AAS68608	Aas68608 DNA encod
2	780	71.6	1337	10 ADD71032	Add71032 Human ald
3	773	70.9	1549	12 ADK70274	Adk70274 Respirato
4	729	66.9	1508	3 AAC98140	Aac98140 Human col
5	729	66.9	1560	12 ADJ75119	Adj75119 Marker ge
6	729	66.9	1560	12 ADN04246	Adn04246 Antipsori
7	729	66.9	1560	13 ACN38728	Acn38728 Tumour-as
8	729	66.9	1560	13 ADS85007	Ads85007 Human ato
9	664	60.9	1316	5 AAF68405	Aaf68405 Human lun
10	664	60.9	1316	6 ABK38316	Abk38316 cDNA enco

us-10-653-681b-1.Oligo.rng										
	11	664	60.9	1316	7	ADS73134	Ads73134	Human	kid	
	12	664	60.9	1316	8	ACA10645	Aca10645	Human	lun	
	13	664	60.9	1316	8	ABX99596	Abx99596	Lung	canc	
	14	664	60.9	1316	10	ADH45842	Adh45842	Human	lun	
	15	664	60.9	1316	12	ADE72379	Ade72379	Human	lun	
	16	664	60.9	1316	13	ADJ19761	Adj19761	Human	lun	
	17	529	48.5	770	13	ADR98739	Adr98739	Lung	spec	
	18	529	48.5	1621	12	ADH13722	Adh13722	Human	ENZ	
c	19	389	35.7	558	10	ABZ84625	Abz84625	Toxicolog		
	20	328	30.1	364	4	AAS39335	Aas39335	Novel	hum	
	21	316	29.0	585	2	AAZ24592	Aaz24592	Human	lun	
	22	316	29.0	585	3	AAC65831	Aac65831	Human	lun	
	23	316	29.0	585	6	ABL49050	Ab149050	Human	lun	
	24	316	29.0	585	6	ABQ92236	Abq92236	Human	lun	
	25	316	29.0	585	9	ADA28651	Ada28651	Human	lun	
	26	316	29.0	585	10	ADE53611	Ade53611	Human	lun	
	27	316	29.0	585	10	ADH36746	Adh36746	Human	lun	
	28	316	29.0	585	12	ADM56549	Adm56549	Human	lun	
	29	316	29.0	585	12	ADN89593	Adn89593	Human	lun	
c	30	316	29.0	857	9	ADA28650	Ada28650	Human	lun	
c	31	316	29.0	858	2	AAZ24591	Aaz24591	Human	lun	
c	32	316	29.0	858	3	AAC65830	Aac65830	Human	lun	
c	33	316	29.0	858	6	ABL49049	Ab149049	Human	lun	
c	34	316	29.0	858	6	ABQ92235	Abq92235	Human	lun	
c	35	316	29.0	858	10	ADE53610	Ade53610	Human	lun	
c	36	316	29.0	858	10	ADH36745	Adh36745	Human	lun	
c	37	316	29.0	858	12	ADM56548	Adm56548	Human	lun	
c	38	316	29.0	858	12	ADN89592	Adn89592	Human	lun	
	39	304	27.9	356	4	AAS39333	Aas39333	Novel	hum	
	40	252	23.1	1396	5	AAS91091	Aas91091	DNA	encod	
	41	249	22.8	861	5	AAS68606	Aas68606	DNA	encod	
	42	235	21.6	830	4	AAI92428	Aai92428	Human	pol	
	43	183	16.8	540	12	ADP28822	Adp28822	Human	sec	
	44	174	16.0	857	13	ADR98738	Adr98738	Lung	spec	
c	45	159	14.6	198	12	ACH84956	Ach84956	Human	gen	

ALIGNMENTS

RESULT 1

AAS68608

ID AAS68608 standard; cDNA; 1337 BP.

XX

AC AAS68608;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #4412.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

Sequence 1337 BP; 390 A; 305 C; 318 G; 324 T; 0 U; 0 other:

Query Match 71.6%; Score 780; DB 5; Length 1337;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 780; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Page 3

us-10-653-681b-1.Oligo.rng

Qy	611	GTCTGTGACACCAGCACGCATTGTTGAGAACATTTCAGGTCTTTGACTTTAAATTGAGTGA	670
Db	858	GTCTGTGACACCAGCACGCATTGTTGAGAACATTTCAGGTCTTTGACTTTAAATTGAGTGA	917
Qy	671	TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA	730
Db	918	TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA	977
Qy	731	ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG	790
Db	978	ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG	1037
Qy	791	AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	850
Db	1038	AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	1097
Qy	851	TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTAGTCCTGTTATAGACGAGAATCGAGG	910
Db	1098	TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTAGTCCTGTTATAGACGAGAATCGAGG	1157
Qy	911	TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA	970
Db	1158	TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA	1217
Qy	971	TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	1030
Db	1218	TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	1277
Qy	1031	AGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAAATAATCAT	1090
Db	1278	AGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAAATAATCAT	1337

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 11:53:34 ; Search time 4250 Seconds
(without alignments)
9762.366 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaataataatcat 1090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result			%		Query		DB	ID	Description
	No.	Score	Match	Length	Length	Match			
c	1	629	57.7	735	5	BM981698		BM981698	UI-CF-EN1
	2	585	53.7	613	6	CB132708		CB132708	K-EST0183
c	3	570	52.3	593	5	BU677104		BU677104	UI-CF-DU1
	4	517	47.4	588	2	BE785963		BE785963	601478213
c	5	496	45.5	540	1	AA804597		AA804597	nk97e06.s
	6	493	45.2	623	4	BM793014		BM793014	K-EST0073
c	7	479	43.9	595	5	BM983180		BM983180	UI-CF-EN1
c	8	454	41.7	644	6	CA450136		CA450136	UI-CF-FN0

c	9	446	40.9	456	1	AI292337	AI292337	qm77c02.x
c	10	446	40.9	458	1	AI744504	AI744504	wg09a09.x
c	11	446	40.9	460	1	AI291463	AI291463	qm73h04.x
c	12	444	40.7	452	1	AI393702	AI393702	tg66d01.x
c	13	432	39.6	446	5	BX104876	BX104876	BX104876
	14	431	39.5	592	4	BM819663	BM819663	K-EST0087
c	15	426	39.1	448	5	BM975664	BM975664	UI-CF-EN1
	16	417	38.3	445	6	CB161124	CB161124	K-EST0220
	17	401	36.8	635	7	CV334625	CV334625	IL3-UT011
	18	400	36.7	620	6	CB118695	CB118695	K-EST0165
c	19	392	36.0	746	2	BF688991	BF688991	602185236
	20	385	35.3	453	4	BG197874	BG197874	RST17122
c	21	378	34.7	388	1	AI831519	AI831519	wj49h11.x
c	22	363	33.3	445	1	AI301329	AI301329	qn27e09.x
c	23	363	33.3	496	2	AW379341	AW379341	MRO-HT024
	24	361	33.1	384	2	BE787870	BE787870	601479812
	25	360	33.0	796	4	BG682196	BG682196	602629503
c	26	345	31.7	388	2	BE711936	BE711936	QV2-HT069
	27	341	31.3	485	7	CV334678	CV334678	IL3-UT011
c	28	338	31.0	415	1	AA947514	AA947514	oq53h01.s
	29	329	30.2	786	5	BQ221381	BQ221381	AGENCOURT
	30	324	29.7	365	5	BQ377471	BQ377471	IL2-UM007
	31	319	29.3	319	4	BM855292	BM855292	K-EST0138
	32	316	29.0	729	5	BX480365	BX480365	DKFZp6860
	33	316	29.0	1177	4	BM558121	BM558121	AGENCOURT
c	34	314	28.8	666	1	AI924753	AI924753	wn58a02.x
c	35	300	27.5	300	1	AI270664	AI270664	qu90c04.x
c	36	299	27.4	549	5	BQ315282	BQ315282	RC3-IT001
	37	299	27.4	909	4	BG169378	BG169378	602320937
	38	296	27.2	316	5	BU175462	BU175462	AGENCOURT
	39	295	27.1	310	4	BM783186	BM783186	K-EST0061
c	40	292	26.8	346	2	BE775022	BE775022	IL2-UM007
	41	285	26.1	582	5	BP278752	BP278752	BP278752
	42	280	25.7	581	5	BP263763	BP263763	BP263763
	43	272	25.0	912	5	BQ220848	BQ220848	AGENCOURT
	44	248	22.8	535	6	CB147729	CB147729	K-EST0203
	45	248	22.8	557	4	BG490449	BG490449	602519494

ALIGNMENTS

RESULT 1

BM981698/c

LOCUS BM981698 735 bp mRNA linear EST 21-FEB-2003

DEFINITION UI-CF-EN1-adi-f-17-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone
UI-CF-EN1-adi-f-17-0-UI 3', mRNA sequence.

ACCESSION BM981698

VERSION BM981698.1 GI:19604453

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 735)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 1-44, >POLY_A#Simple_repeat (matched compliment)
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES Location/Qualifiers
source 1. .735
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-adi-f-17-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-EN1 is a normalized cDNA library containing the
following tissue(s): Primary Lung Cystic Fibrosis
Epithelial Cells. The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT7T3-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is CTGCTCAGGT.
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_LIB=UI-CF-EN1
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 57.7%; Score 629; DB 5; Length 735;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	462	GGCTCTCCGGATAGACCTTGGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAG	521
Db	645	GGCTCTCCGGATAGACCTTGGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAG	586
Qy	522	ATTAAGGAGATTGCTGCAAAGCACAAAAAACCGCAGCCCAGGTTCTGATCCGTTTCCAT	581
Db	585	ATTAAGGAGATTGCTGCAAAGCACAAAAAACCGCAGCCCAGGTTCTGATCCGTTTCCAT	526
Qy	582	ATCCAGAGGAATGTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAAC	641
Db	525	ATCCAGAGGAATGTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAAC	466
Qy	642	ATTCAGGTCTTTGACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAAC	701
Db	465	ATTCAGGTCTTTGACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAAC	406
Qy	702	AGAAACTGGAGGGCCTGTAACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGAT	761
Db	405	AGAAACTGGAGGGCCTGTAACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGAT	346
Qy	762	GCAGAAATATTGAGGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTG	821
Db	345	GCAGAAATATTGAGGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTG	286
Qy	822	AAGTGTGACTACCTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAA	881
Db	285	AAGTGTGACTACCTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAA	226
Qy	882	CTTAGTCCTGTTATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTT	941
Db	225	CTTAGTCCTGTTATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTT	166
Qy	942	AACTAGGATCAGAATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCC	1001
Db	165	AACTAGGATCAGAATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCC	106
Qy	1002	ACTTATCTGATCAGAACAAATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATG	1061
Db	105	ACTTATCTGATCAGAACAAATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATG	46
Qy	1062	TAAAGATCAATAAAAAAAAAATAAATCAT	1090
Db	45	TAAAGATCAATAAAAAAAAAATAAATCAT	17

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 06:02:48 ; Search time 5146 Seconds
(without alignments)
10263.542 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	833	76.4	1337	9	HSU37100	U37100 Homo sapien
2	830.8	76.2	1336	6	CQ718316	CQ718316 Sequence

3	828.2	76.0	1551	9	BC008837	BC008837 Homo sapi
4	828.2	76.0	1560	6	CQ776685	CQ776685 Sequence
5	820.2	75.2	1611	9	AF524864	AF524864 Homo sapi
6	796.6	73.1	1316	6	AR272611	AR272611 Sequence
7	796.6	73.1	1316	6	AR276192	AR276192 Sequence
8	796.6	73.1	1316	6	AR406467	AR406467 Sequence
9	796.6	73.1	1316	6	AR440317	AR440317 Sequence
10	796.6	73.1	1316	6	AR472475	AR472475 Sequence
11	796.6	73.1	1316	6	AR543128	AR543128 Sequence
12	796.6	73.1	1316	6	AX062696	AX062696 Sequence
13	796.6	73.1	1316	6	AX367613	AX367613 Sequence
14	796.6	73.1	1316	9	AF052577	AF052577 Homo sapi
15	718	65.9	1315	6	AX743782	AX743782 Sequence
16	651.8	59.8	3994	9	AL669847	AL669847 Human DNA
17	651.8	59.8	121210	9	AL607022	AL607022 Human DNA
18	574	52.7	574	9	AF044961	AF044961 Homo sapi
19	508	46.6	1080	6	AX772965	AX772965 Sequence
20	459.4	42.1	951	9	BT006794	BT006794 Homo sapi
21	459.4	42.1	951	12	BT007750	BT007750 Synthetic
22	458.4	42.1	948	9	CR541801	CR541801 Homo sapi
23	431	39.5	951	6	AX380448	AX380448 Sequence
24	431	39.5	951	6	AX772962	AX772962 Sequence
25	395.6	36.3	473	9	AY347931	AY347931 Macaca ra
26	377.8	34.7	144234	2	AP002425	AP002425 Homo sapi
c 27	377.8	34.7	144279	2	AP001570	AP001570 Homo sapi
c 28	377.8	34.7	216972	9	AC067819	AC067819 Homo sapi
29	374.8	34.4	585	6	CQ732993	CQ732993 Sequence
c 30	361	33.1	137557	9	AC005909	AC005909 Homo sapi
31	360.4	33.1	364	6	AX247463	AX247463 Sequence
32	358.4	32.9	163631	9	AC009276	AC009276 Homo sapi
c 33	358.4	32.9	170919	9	AC078847	AC078847 Homo sapi
c 34	358.4	32.9	177373	2	AP002452	AP002452 Homo sapi
35	358.4	32.9	196039	2	AC055757	AC055757 Homo sapi
36	357.2	32.8	1268	10	CGU81045	U81045 Cricetulus
37	354.4	32.5	1400	10	BC037690	BC037690 Mus muscu
38	353.4	32.4	356	6	AX247461	AX247461 Sequence
39	353.4	32.4	1446	10	BC079133	BC079133 Rattus no
40	344.6	31.6	1413	10	AF182168	AF182168 Rattus no
41	332	30.5	1315	10	BC005789	BC005789 Mus muscu
42	331.6	30.4	1304	10	MMU04204	U04204 Mus musculu
43	327.2	30.0	1225	6	CQ777549	CQ777549 Sequence
44	327.2	30.0	1225	10	MUSMVP	J05663 Mouse vas d
45	327	30.0	993	10	RNO277957	AJ277957 Rattus no

ALIGNMENTS

RESULT 1
 HSU37100
 LOCUS HSU37100 1337 bp mRNA linear PRI 28-MAY-1998
 DEFINITION Homo sapiens aldose reductase-like peptide mRNA, complete cds.
 ACCESSION U37100
 VERSION U37100.1 GI:3150034
 KEYWORDS .

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1337)
 AUTHORS Cao,D., Fan,S.T. and Chung,S.S.
 TITLE Identification and characterization of a novel human aldose
 reductase-like gene
 JOURNAL J. Biol. Chem. 273 (19), 11429-11435 (1998)
 MEDLINE 98234319
 PUBMED 9565553
 REFERENCE 2 (bases 1 to 1337)
 AUTHORS Cao,D.
 TITLE Direct Submission
 JOURNAL Submitted (27-SEP-1995) Deliang Cao, The University of Hong Kong,
 Institute of Molecular Biology, 8 Sassoon Road, Pokfulam, Hong
 Kong, Hong Kong

FEATURES Location/Qualifiers
 source 1. .1337
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="pCDL-1"
 /tissue_type="liver tumor"
 /dev_stage="adult"
 CDS 70. .1020
 /codon_start=1
 /product="aldose reductase-like peptide"
 /protein_id="AAC17469.1"
 /db_xref="GI:3150035"

/translation="MATFVELSTKAKMPIVGLGTWKSPLGKVKEAVKVAIDAGYRHID
 CAYVYQNEHEVGEAIQEKIQEKAVKREDLFIVSKLWPTFFERPLVRKAFEKTLKDLKL
 SYLDVYLIHWPQGFKSGDDLFPKDDKGNAIGGKATFLDAWEAMEELVDEGLVKALGVS
 NFSHFQIEKLLNKPGLKYKPVTNQVECHPYLTQEKLIQYCHSKGITVTAYSPLGSPDR
 PWAKPEDPSLLEDPKIKEIAAKHKKTAAQVLIRFHIQRNVIVIPKSVTPARIVENIQV
 FDFKLSDEEMATILSFNRNWRACNVLQSSHLEDYPFDAEY"
 polyA_site 1317. .1322

ORIGIN

Query Match 76.4%; Score 833; DB 9; Length 1337;
 Best Local Similarity 81.5%; Pred. No. 1.4e-221;
 Matches 1090; Conservative 0; Mismatches 0; Indels 247; Gaps

1;

Qy	1	CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT	60
Db	1	CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT	60
Qy	61	GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC	120
Db	61	GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC	120

Qy	121	CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT	180
Db	121	CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT	180
Qy	181	GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA	240
Db	181	GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA	240
Qy	241	GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC	300
Db	241	GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC	300
Qy	301	AAGTTGTGGC-----	310
Db	301	AAGTTGTGGCCACTTTCTTTGAGAGACCCCTTGTGAGGAAAGCCTTTGAGAAGACCCCTC	360
Qy	311	-----	310
Db	361	AAGGACCTGAAGCTGAGCTATCTGGACGTCTATCTTATTCACTGGCCACAGGGATTCAAG	420
Qy	311	-----	310
Db	421	TCTGGGGATGACCTTTTCCCAAAGATGATAAAGGTAATGCCATCGGTGGAAAAGCAACG	480
Qy	311	-----	310
Db	481	TTCTTGGATGCCTGGGAGGCCATGGAGGAGCTGGTGGATGAGGGGCTGGTGAAGCCCTT	540
Qy	311	-----CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAA	353
Db	541	GGGGTCTCCAATTTTCAGCCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAA	600
Qy	354	TATAAACCAGTGACTAACCAGGTTGAGTGTACCCCATACCTCACGCAGGAGAAACTGATC	413
Db	601	TATAAACCAGTGACTAACCAGGTTGAGTGTACCCCATACCTCACGCAGGAGAAACTGATC	660
Qy	414	CAGTACTGCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGAT	473
Db	661	CAGTACTGCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGAT	720
Qy	474	AGACCTTGGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATT	533
Db	721	AGACCTTGGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATT	780
Qy	534	GCTGCAAAGCACAAAAAACCAGCCAGGTTCTGATCCGTTTCCATATCCAGAGGAAT	593
Db	781	GCTGCAAAGCACAAAAAACCAGCCAGGTTCTGATCCGTTTCCATATCCAGAGGAAT	840
Qy	594	GTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTACAGGTCTTT	653
Db	841	GTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTACAGGTCTTT	900
Qy	654	GACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGG	713
Db	901	GACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGG	960

Qy	714	GCCTGTAACGTGTTGCAATCCTCTCATTGGAAGACTATCCCTTCGATGCAGAATATTGA	773
Db	961	GCCTGTAACGTGTTGCAATCCTCTCATTGGAAGACTATCCCTTCGATGCAGAATATTGA	
1020			
Qy	774	GGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTCTTCGCTGAAGTGTGACTAC	833
Db	1021	GGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTCTTCGCTGAAGTGTGACTAC	
1080			
Qy	834	CTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTT	893
Db	1081	CTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTT	
1140			
Qy	894	ATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAG	953
Db	1141	ATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAG	
1200			
Qy	954	AATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATC	
1013			
Db	1201	AATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATC	
1260			
Qy	1014	AGAACAAATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATA	
1073			
Db	1261	AGAACAAATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATA	
1320			
Qy	1074	AAAAAAATAATAATCAT	1090
Db	1321	AAAAAAATAATAATCAT	1337

us-10-653-681b-1.rng

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 05:31:49 ; Search time 697 Seconds
(without alignments)
9257.558 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	833	76.4	1337	5	AAS68608	Aas68608 DNA encod
2	833	76.4	1337	10	ADD71032	Add71032 Human ald
3	832.2	76.3	1508	3	AAC98140	Aac98140 Human col
4	828.2	76.0	1560	12	ADJ75119	Adj75119 Marker ge
5	828.2	76.0	1560	12	ADN04246	Adn04246 Antipsori
6	828.2	76.0	1560	13	ACN38728	Acn38728 Tumour-as
7	828.2	76.0	1560	13	ADS85007	Ads85007 Human ato
8	822.8	75.5	1549	12	ADK70274	Adk70274 Respirato
9	796.6	73.1	1316	5	AAF68405	Aaf68405 Human lun
10	796.6	73.1	1316	6	ABK38316	Abk38316 cDNA enco

us-10-653-681b-1.rng							
11	796.6	73.1	1316	7	ADS73134	Ads73134	Human kid
12	796.6	73.1	1316	8	ACA10645	Aca10645	Human lun
13	796.6	73.1	1316	8	ABX99596	Abx99596	Lung canc
14	796.6	73.1	1316	10	ADH45842	Adh45842	Human lun
15	796.6	73.1	1316	12	ADE72379	Ade72379	Human lun
16	796.6	73.1	1316	13	ADJ19761	Adj19761	Human lun
17	770.8	70.7	1621	12	ADH13722	Adh13722	Human ENZ
18	718	65.9	1315	10	ADC97771	Adc97771	Human ARL
19	635.4	58.3	1816	11	ACN92921	Acn92921	Breast ca
20	616	56.5	770	13	ADR98739	Adr98739	Lung spec
21	540.2	49.6	1170	12	ADH45334	Adh45334	Human enz
22	508	46.6	1080	9	ACC83986	Acc83986	Human ald
C 23	439.8	40.3	558	10	ABZ84625	Abz84625	Toxicolog
24	432	39.6	971	10	ADC10183	Adc10183	Human NOV
25	431	39.5	951	6	ABA94733	Aba94733	Human dru
26	431	39.5	966	10	ADC10185	Adc10185	Human NOV
27	360.4	33.1	364	4	AAS39335	Aas39335	Novel hum
28	353.4	32.4	356	4	AAS39333	Aas39333	Novel hum
29	333.8	30.6	1926	5	AAS72230	Aas72230	DNA encod
30	333.8	30.6	1926	5	AAS92672	Aas92672	DNA encod
31	333.8	30.6	3620	5	AAS69995	Aas69995	DNA encod
32	327.2	30.0	1225	12	ADJ75983	Adj75983	Marker ge
33	317	29.1	540	12	ADP28822	Adp28822	Human sec
34	316	29.0	585	2	AAZ24592	Aaz24592	Human lun
35	316	29.0	585	3	AAC65831	Aac65831	Human lun
36	316	29.0	585	6	ABL49050	Ab149050	Human lun
37	316	29.0	585	6	ABQ92236	Abq92236	Human lun
38	316	29.0	585	9	ADA28651	Ada28651	Human lun
39	316	29.0	585	10	ADE53611	Ade53611	Human lun
40	316	29.0	585	10	ADH36746	Adh36746	Human lun
41	316	29.0	585	12	ADM56549	Adm56549	Human lun
42	316	29.0	585	12	ADN89593	Adn89593	Human lun
C 43	316	29.0	857	9	ADA28650	Ada28650	Human lun
C 44	316	29.0	858	2	AAZ24591	Aaz24591	Human lun
C 45	316	29.0	858	3	AAC65830	Aac65830	Human lun

ALIGNMENTS

RESULT 1

AAS68608

ID AAS68608 standard; cDNA; 1337 BP.

XX

AC AAS68608;

XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #4412.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG04421.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 4412; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1337 BP; 390 A; 305 C; 318 G; 324 T; 0 U; 0 Other;

Query Match 76.4%; Score 833; DB 5; Length 1337;
Best Local Similarity 81.5%; Pred. No. 1.1e-230;
Matches 1090; Conservative 0; Mismatches 0; Indels 247; Gaps 1;

Qy	1	CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT	60
Db	1	CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT	60
Qy	61	GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC	120
Db	61	GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC	120
Qy	121	CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT	180
Db	121	CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT	180
Qy	181	GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA	240
Db	181	GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA	240
Qy	241	GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC	300
Db	241	GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC	300

us-10-653-681b-1.rng

Qy 301 AAGTTGTGGC----- 310
 Db 301 AAGTTGTGGCCCACTTTCTTTGAGAGACCCCTTGTGAGGAAAGCCTTTGAGAAGACCCTC 360
 Qy 311 ----- 310
 Db 361 AAGGACCTGAAGCTGAGCTATCTGGACGTCTATCTTATTCACTGGCCACAGGGATTCAAG 420
 Qy 311 ----- 310
 Db 421 TCTGGGGATGACCTTTTCCCAAAGATGATAAAGGTAATGCCATCGGTGAAAAAGCAACG 480
 Qy 311 ----- 310
 Db 481 TTCTTGGATGCCTGGGAGGCCATGGAGGAGCTGGTGGATGAGGGGCTGGTGAAAGCCCTT 540
 Qy 311 -----CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAA 353
 Db 541 GGGGTCTCCAATTTTCAGCCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAA 600
 Qy 354 TATAAACCAGTGACTAACCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATC 413
 Db 601 TATAAACCAGTGACTAACCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATC 660
 Qy 414 CAGTACTGCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGAT 473
 Db 661 CAGTACTGCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGAT 720
 Qy 474 AGACCTTGGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATT 533
 Db 721 AGACCTTGGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATT 780
 Qy 534 GCTGCAAAGCACAAAAAACCGCAGCCAGGTTCTGATCCGTTTCCATATCCAGAGGAAT 593
 Db 781 GCTGCAAAGCACAAAAAACCGCAGCCAGGTTCTGATCCGTTTCCATATCCAGAGGAAT 840
 Qy 594 GTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTTCAGGTCTTT 653
 Db 841 GTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTTCAGGTCTTT 900
 Qy 654 GACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGG 713
 Db 901 GACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGG 960
 Qy 714 GCCTGTAACGTGTTGCAATCCTCTCATTGGAAGACTATCCCTTCGATGCAGAATATTGA 773
 Db 961 GCCTGTAACGTGTTGCAATCCTCTCATTGGAAGACTATCCCTTCGATGCAGAATATTGA 1020
 Qy 774 GGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTAC 833
 Db 1021 GGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTAC 1080
 Qy 834 CTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTT 893
 Db 1081 CTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTT 1140
 Qy 894 ATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAG 953
 Db 1141 ATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAG 1200
 Qy 954 AATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATC 1013
 Db 1201 AATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATC 1260

us-10-653-681b-1.rng

```
Qy      1014 AGAACAAATGTTTATTAAGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATA 1073
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1261 AGAACAAATGTTTATTAAGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATA 1320
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1074 AAAAAAATAATAATCAT 1090
         ||||||||||||||||
Db      1321 AAAAAAATAATAATCAT 1337
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 07:25:59 ; Search time 4248 Seconds
(without alignments)
9766.962 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	
c 1	705.4	64.7	735	5	BM981698	BM981698 UI-CF-EN1	
2	685.2	62.9	1586	3	CR607509	CR607509 full-leng	
c 3	666	61.1	746	2	BF688991	BF688991 602185236	
c 4	652.2	59.8	796	5	BX337598	BX337598 BX337598	
5	647	59.4	909	4	BG169378	BG169378 602320937	
c 6	586.6	53.8	666	1	AI924753	AI924753 wn58a02.x	
7	585	53.7	613	6	CB132708	CB132708 K-EST0183	
c 8	573	52.6	595	5	BM983180	BM983180 UI-CF-EN1	

c	9	570.8	52.4	593	5	BU677104	BU677104	UI-CF-DU1
	10	542.4	49.8	623	4	BM793014	BM793014	K-EST0073
	11	526.2	48.3	969	7	CO581979	CO581979	ILLUMIGEN
	12	522	47.9	1012	5	BQ943650	BQ943650	AGENCOURT
c	13	521.6	47.9	540	1	AA804597	AA804597	nk97e06.s
c	14	518.2	47.5	644	6	CA450136	CA450136	UI-CF-FN0
	15	517	47.4	588	2	BE785963	BE785963	601478213
	16	506.8	46.5	881	7	CO580792	CO580792	ILLUMIGEN
	17	502.6	46.1	704	7	CO582646	CO582646	ILLUMIGEN
	18	498.8	45.8	620	6	CB118695	CB118695	K-EST0165
	19	481.4	44.2	635	7	CV334625	CV334625	IL3-UT011
c	20	475.4	43.6	496	2	AW379341	AW379341	MRO-HT024
c	21	460	42.2	920	2	BE964368	BE964368	601658069
	22	456.4	41.9	592	4	BM819663	BM819663	K-EST0087
	23	449.4	41.2	786	5	BQ221381	BQ221381	AGENCOURT
c	24	447.2	41.0	452	1	AI393702	AI393702	tg66d01.x
c	25	446.4	41.0	456	1	AI292337	AI292337	qm77c02.x
c	26	446.4	41.0	458	1	AI744504	AI744504	wg09a09.x
c	27	446.4	41.0	460	1	AI291463	AI291463	qm73h04.x
	28	446.4	41.0	582	5	BP278752	BP278752	BP278752
c	29	432	39.6	446	5	BX104876	BX104876	BX104876
c	30	426.6	39.1	445	1	AI301329	AI301329	qn27e09.x
	31	426.4	39.1	914	7	CO775128	CO775128	ILLUMIGEN
c	32	426	39.1	448	5	BM975664	BM975664	UI-CF-EN1
	33	417	38.3	445	6	CB161124	CB161124	K-EST0220
	34	402.4	36.9	581	5	BP263763	BP263763	BP263763
c	35	392.6	36.0	415	1	AA947514	AA947514	oq53h01.s
	36	388.8	35.7	453	4	BG197874	BG197874	RST17122
	37	384.8	35.3	912	5	BQ220848	BQ220848	AGENCOURT
c	38	383.4	35.2	405	6	C75075	C75075	C75075 Huma
c	39	378	34.7	388	1	AI831519	AI831519	wj49h11.x
	40	373	34.2	384	2	BE787870	BE787870	601479812
c	41	370.6	34.0	386	1	AI813308	AI813308	wj33c01.x
c	42	365.2	33.5	388	2	BE711936	BE711936	QV2-HT069
	43	360	33.0	796	4	BG682196	BG682196	602629503
	44	351.2	32.2	1342	3	AK075865	AK075865	Mus muscu
	45	348	31.9	1236	3	AK019906	AK019906	Mus muscu

us-10-653-681b-1.Oligo.rni

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 08:31:19 ; Search time 227 Seconds
(without alignments)
7857.012 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaataataatcat 1090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	729	66.9	1515	4	US-09-949-016-1344	Sequence 1344, Ap
2	664	60.9	1316	4	US-09-702-705-323	Sequence 323, App
3	664	60.9	1316	4	US-09-736-457-323	Sequence 323, App
4	664	60.9	1316	4	US-09-614-124B-323	Sequence 323, App
5	664	60.9	1316	4	US-09-671-325-323	Sequence 323, App
6	664	60.9	1316	4	US-09-589-184-323	Sequence 323, App
7	664	60.9	1316	4	US-09-658-824-323	Sequence 323, App
8	331	30.4	17740	4	US-09-949-016-13086	Sequence 13086, A
9	316	29.0	585	3	US-09-123-912-92	Sequence 92, Appl
10	316	29.0	585	3	US-09-643-597-92	Sequence 92, Appl
11	316	29.0	585	4	US-09-480-884A-92	Sequence 92, Appl
12	316	29.0	585	4	US-09-542-615A-92	Sequence 92, Appl
13	316	29.0	585	4	US-09-606-421B-92	Sequence 92, Appl
14	316	29.0	585	4	US-09-221-107-92	Sequence 92, Appl
15	316	29.0	585	4	US-09-466-396A-92	Sequence 92, Appl
16	316	29.0	585	4	US-09-476-496A-92	Sequence 92, Appl
17	316	29.0	585	4	US-09-630-940B-92	Sequence 92, Appl

					us-10-653-681b-1.oligo.rni	
	18	316	29.0	585	4	US-09-285-479-92 Sequence 92, Appl
c	19	316	29.0	858	3	US-09-123-912-91 Sequence 91, Appl
c	20	316	29.0	858	3	US-09-643-597-91 Sequence 91, Appl
c	21	316	29.0	858	4	US-09-480-884A-91 Sequence 91, Appl
c	22	316	29.0	858	4	US-09-542-615A-91 Sequence 91, Appl
c	23	316	29.0	858	4	US-09-606-421B-91 Sequence 91, Appl
c	24	316	29.0	858	4	US-09-221-107-91 Sequence 91, Appl
c	25	316	29.0	858	4	US-09-466-396A-91 Sequence 91, Appl
c	26	316	29.0	858	4	US-09-476-496A-91 Sequence 91, Appl
c	27	316	29.0	858	4	US-09-630-940B-91 Sequence 91, Appl
c	28	316	29.0	858	4	US-09-285-479-91 Sequence 91, Appl
	29	300	27.5	601	4	US-09-949-016-46452 Sequence 46452, A
	30	261	23.9	601	4	US-09-949-016-46451 Sequence 46451, A
	31	223	20.5	914	4	US-09-949-016-3127 Sequence 3127, Ap
	32	159	14.6	15141	4	US-09-949-016-14869 Sequence 14869, A
	33	156	14.3	601	4	US-09-949-016-46442 Sequence 46442, A
	34	156	14.3	601	4	US-09-949-016-113475 Sequence 113475,
	35	144	13.2	233	4	US-09-702-705-31 Sequence 31, Appl
	36	144	13.2	233	4	US-09-736-457-31 Sequence 31, Appl
	37	144	13.2	233	4	US-09-614-124B-31 Sequence 31, Appl
	38	144	13.2	233	4	US-09-671-325-31 Sequence 31, Appl
	39	144	13.2	233	4	US-09-589-184-31 Sequence 31, Appl
	40	144	13.2	233	4	US-09-658-824-31 Sequence 31, Appl
	41	118	10.8	601	4	US-09-949-016-46437 Sequence 46437, A
	42	118	10.8	601	4	US-09-949-016-113470 Sequence 113470,
	43	26	2.4	1337	3	US-08-801-344-3 Sequence 3, Appli
	44	26	2.4	1337	3	US-09-498-599-3 Sequence 3, Appli
	45	22	2.0	292	4	US-09-313-294A-6562 Sequence 6562, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-1344

; Sequence 1344, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 1344

; LENGTH: 1515

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-1344

Query Match 66.9%; Score 729; DB 4; Length 1515;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 779; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 311 CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACAGTGACTAA 370

us-10-653-681b-1.Oligo.rni

Db	736		CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAACCACTGACTAA	795
Qy	371		CCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATCCAGTACTGCCACTCCAA	430
Db	796		CCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATCCAGTACTGCCACTCCAA	855
Qy	431		GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC	490
Db	856		GGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTTGGGCCAAGCC	915
Qy	491		AGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA	550
Db	916		AGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAAAGCACAAAAA	975
Qy	551		AACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA	610
Db	976		AACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTGTCATCCCCAA	1035
Qy	611		GTCTGTGACACCAGCACGCATTGTTGAGAACATTACAGTCTTTGACTTTAAATTGAGTGA	670
Db	1036		GTCTGTGACACCAGCACGCATTGTTGAGAACATTACAGTCTTTGACTTTAAATTGAGTGA	1095
Qy	671		TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA	730
Db	1096		TGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTAACGTGTTGCA	1155
Qy	731		ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG	790
Db	1156		ATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAATCTCCTGGTG	1215
Qy	791		AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	850
Db	1216		AGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACTCATGTCCCAT	1275
Qy	851		TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTAGTCCTGTTATAGACGAGAATCGAGG	910
Db	1276		TTTAGCCAAGCTTATTTAAGATCACAGTGAACCTAGTCCTGTTATAGACGAGAATCGAGG	1335
Qy	911		TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA	970
Db	1336		TGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCACAGAAAAGCA	1395
Qy	971		TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	1030
Db	1396		TGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAAATGTTTATTA	1455
Qy	1031		AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAATAATAATCAT	1090
Db	1456		AGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAATAATAATCAT	1515

us-10-653-681b-1.Oligo.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 08:43:09 ; Search time 839 Seconds
(without alignments)
8421.615 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 7297361 seqs, 3241162794 residues

Word size : 0

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

us-10-653-681b-1.oligo.rnpb

Result No.	Score	Query Match	Length	DB	ID	Description
1	1090	100.0	1090	21	US-10-653-681A-1	Sequence 1, Appli
2	729	66.9	1508	9	US-09-925-299-150	Sequence 150, App
3	729	66.9	1508	10	US-09-925-299-150	Sequence 150, App
4	664	60.9	1316	9	US-09-736-457-323	Sequence 323, App
5	664	60.9	1316	9	US-09-902-941-323	Sequence 323, App
6	664	60.9	1316	9	US-09-849-626-323	Sequence 323, App
7	664	60.9	1316	10	US-09-476-300-323	Sequence 323, App
8	664	60.9	1316	14	US-10-017-754-323	Sequence 323, App
9	664	60.9	1316	15	US-10-102-524-1731	Sequence 1731, Ap
10	664	60.9	1316	16	US-10-113-872-323	Sequence 323, App
11	664	60.9	1316	17	US-10-283-017-323	Sequence 323, App
12	487	44.7	1279	21	US-10-653-681A-3	Sequence 3, Appli
13	328	30.1	364	10	US-09-803-719-2393	Sequence 2393, Ap
14	316	29.0	585	9	US-09-735-705-92	Sequence 92, Appl
15	316	29.0	585	9	US-09-850-716A-92	Sequence 92, Appl
16	316	29.0	585	9	US-09-897-778-92	Sequence 92, Appl
17	316	29.0	585	10	US-09-466-396A-92	Sequence 92, Appl
18	316	29.0	585	14	US-10-007-700-92	Sequence 92, Appl
19	316	29.0	585	15	US-10-117-982-92	Sequence 92, Appl
20	316	29.0	585	17	US-10-313-986-92	Sequence 92, Appl
21	316	29.0	585	20	US-10-775-972-92	Sequence 92, Appl
22	316	29.0	585	22	US-10-922-124-92	Sequence 92, Appl
C 23	316	29.0	858	9	US-09-735-705-91	Sequence 91, Appl
C 24	316	29.0	858	9	US-09-850-716A-91	Sequence 91, Appl
C 25	316	29.0	858	9	US-09-897-778-91	Sequence 91, Appl
C 26	316	29.0	858	10	US-09-466-396A-91	Sequence 91, Appl
C 27	316	29.0	858	14	US-10-007-700-91	Sequence 91, Appl
C 28	316	29.0	858	15	US-10-117-982-91	Sequence 91, Appl
C 29	316	29.0	858	17	US-10-313-986-91	Sequence 91, Appl
C 30	316	29.0	858	20	US-10-775-972-91	Sequence 91, Appl
C 31	316	29.0	858	22	US-10-922-124-91	Sequence 91, Appl
32	304	27.9	356	10	US-09-803-719-2391	Sequence 2391, Ap
C 33	159	14.6	198	16	US-10-029-386-18151	Sequence 18151, A
C 34	159	14.6	546	16	US-10-029-386-4451	Sequence 4451, Ap
35	159	14.6	951	15	US-10-274-694-36	Sequence 36, Appl
36	159	14.6	951	20	US-10-332-448-36	Sequence 36, Appl
37	159	14.6	1315	15	US-10-274-375-1	Sequence 1, Appli
38	144	13.2	233	9	US-09-736-457-31	Sequence 31, Appl
39	144	13.2	233	9	US-09-902-941-31	Sequence 31, Appl
40	144	13.2	233	9	US-09-849-626-31	Sequence 31, Appl
41	144	13.2	233	10	US-09-476-300-31	Sequence 31, Appl
42	144	13.2	233	14	US-10-017-754-31	Sequence 31, Appl
43	144	13.2	233	16	US-10-113-872-31	Sequence 31, Appl
44	144	13.2	233	17	US-10-283-017-31	Sequence 31, Appl
45	116	10.6	1816	14	US-10-198-846-14071	Sequence 14071, A

ALIGNMENTS

RESULT 1

US-10-653-681A-1

; Sequence 1, Application US/10653681A

; Publication No. US20050048503A1

; GENERAL INFORMATION:

; APPLICANT: DAI, KEN-SHOW

; TITLE OF INVENTION: HUMAN ARL-RELATED GENE VARIANTS ASSOCIATED WITH CANCER

; FILE REFERENCE: U014798-3

; CURRENT APPLICATION NUMBER: US/10/653,681A

; CURRENT FILING DATE: 2003-09-02

us-10-653-681b-1.oligo.rnpb

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 1090
;   TYPE: DNA
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: VARIANT OF HUMAN ALDOSE REDUCTASE-LIKE GENE
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (70)..(333)
US-10-653-681A-1

```

Query Match 100.0%; Score 1090; DB 21; Length 1090;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT 60

Qy     61 GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC 120

Qy    121 CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT 180

Qy    181 GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA 240

Qy    241 GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC 300

Qy    301 AAGTTGTGGCCCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAAC 360
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AAGTTGTGGCCCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAATATAAAC 360

Qy    361 CAGTGACTAACCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATCCAGTACT 420
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 CAGTGACTAACCAGGTTGAGTGTACCCATACCTCACGCAGGAGAACTGATCCAGTACT 420

Qy    421 GCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTT 480
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGATAGACCTT 480

Qy    481 GGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAA 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GGGCCAAGCCAGAAGACCCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATTGCTGCAA 540

Qy    541 AGCACAAAAAAACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTG 600
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 AGCACAAAAAAACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAATGTGATTG 600

Qy    601 TCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTCAGGTCTTTGACTTTA 660
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTCAGGTCTTTGACTTTA 660

Qy    661 AATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAAACTGGAGGGCCTGTA 720
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```

db	661	AATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAAACTGGAGGGCCTGTA	720
Qy	721	ACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAA 	780
db	721	ACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAA	780
Qy	781	TCTCCTGGTGAGATTATACAGGAGATTCTCTTCTTCGCTGAAGTGTGACTACCTCCACT 	840
db	781	TCTCCTGGTGAGATTATACAGGAGATTCTCTTCTTCGCTGAAGTGTGACTACCTCCACT	840
Qy	841	CATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACG 	900
db	841	CATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACG	900
Qy	901	AGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCA 	960
db	901	AGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCA	960
Qy	961	CAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAA 	1020
db	961	CAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAA	1020
Qy	1021	ATGTTTTATTAAGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAA 	1080
db	1021	ATGTTTTATTAAGCATCAGAAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAA	1080
Qy	1081	TAATAATCAT 1090 	
db	1081	TAATAATCAT 1090	

us-10-653-681b-1.rni

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 00:34:58 ; Search time 226 Seconds
(without alignments)
7891.777 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	828.2	76.0	1515	4	US-09-949-016-1344 Sequence 1344, Ap
2	796.6	73.1	1316	4	US-09-702-705-323 Sequence 323, App
3	796.6	73.1	1316	4	US-09-736-457-323 Sequence 323, App
4	796.6	73.1	1316	4	US-09-614-124B-323 Sequence 323, App
5	796.6	73.1	1316	4	US-09-671-325-323 Sequence 323, App
6	796.6	73.1	1316	4	US-09-589-184-323 Sequence 323, App
7	796.6	73.1	1316	4	US-09-658-824-323 Sequence 323, App
8	389.2	35.7	914	4	US-09-949-016-3127 Sequence 3127, Ap
9	358.4	32.9	17740	4	US-09-949-016-13086 Sequence 13086, A
10	328.6	30.1	601	4	US-09-949-016-46452 Sequence 46452, A
11	316	29.0	585	3	US-09-123-912-92 Sequence 92, Appl
12	316	29.0	585	3	US-09-643-597-92 Sequence 92, Appl
13	316	29.0	585	4	US-09-480-884A-92 Sequence 92, Appl
14	316	29.0	585	4	US-09-542-615A-92 Sequence 92, Appl
15	316	29.0	585	4	US-09-606-421B-92 Sequence 92, Appl
16	316	29.0	585	4	US-09-221-107-92 Sequence 92, Appl
17	316	29.0	585	4	US-09-466-396A-92 Sequence 92, Appl

					us-10-653-681b-1.rni		
	18	316	29.0	585	4	US-09-476-496A-92	Sequence 92, Appl
	19	316	29.0	585	4	US-09-630-940B-92	Sequence 92, Appl
	20	316	29.0	585	4	US-09-285-479-92	Sequence 92, Appl
C	21	316	29.0	858	3	US-09-123-912-91	Sequence 91, Appl
C	22	316	29.0	858	3	US-09-643-597-91	Sequence 91, Appl
C	23	316	29.0	858	4	US-09-480-884A-91	Sequence 91, Appl
C	24	316	29.0	858	4	US-09-542-615A-91	Sequence 91, Appl
C	25	316	29.0	858	4	US-09-606-421B-91	Sequence 91, Appl
C	26	316	29.0	858	4	US-09-221-107-91	Sequence 91, Appl
C	27	316	29.0	858	4	US-09-466-396A-91	Sequence 91, Appl
C	28	316	29.0	858	4	US-09-476-496A-91	Sequence 91, Appl
C	29	316	29.0	858	4	US-09-630-940B-91	Sequence 91, Appl
C	30	316	29.0	858	4	US-09-285-479-91	Sequence 91, Appl
	31	304.8	28.0	1335	4	US-09-023-655-1010	Sequence 1010, Ap
	32	292	26.8	1337	3	US-08-801-344-3	Sequence 3, Appli
	33	292	26.8	1337	3	US-09-498-599-3	Sequence 3, Appli
	34	288.4	26.5	601	4	US-09-949-016-46451	Sequence 46451, A
	35	170.2	15.6	1290	4	US-09-270-767-13724	Sequence 13724, A
	36	169.8	15.6	15141	4	US-09-949-016-14869	Sequence 14869, A
	37	165.4	15.2	601	4	US-09-949-016-46442	Sequence 46442, A
	38	165.4	15.2	601	4	US-09-949-016-113475	Sequence 113475,
	39	145.4	13.3	233	4	US-09-702-705-31	Sequence 31, Appl
	40	145.4	13.3	233	4	US-09-736-457-31	Sequence 31, Appl
	41	145.4	13.3	233	4	US-09-614-124B-31	Sequence 31, Appl
	42	145.4	13.3	233	4	US-09-671-325-31	Sequence 31, Appl
	43	145.4	13.3	233	4	US-09-589-184-31	Sequence 31, Appl
	44	145.4	13.3	233	4	US-09-658-824-31	Sequence 31, Appl
	45	141.6	13.0	292	4	US-09-313-294A-6562	Sequence 6562, Ap

ALIGNMENTS

RESULT 1

```

US-09-949-016-1344
; Sequence 1344, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1344
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1344

```

Query Match 76.0%; Score 828.2; DB 4; Length 1515;
 Best Local Similarity 81.3%; Pred. No. 1.2e-247;
 Matches 1087; Conservative 0; Mismatches 3; Indels 247; Gaps 1;

QY 1 CAAAAACAGCAACAGAAAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTCT 60

us-10-653-681b-1.rni

Db	179		CAAAAACAGCAACAGAGAGCAGGACGTGAGACTTCTACCTGCTCACTCAGAATCATTTCT	238
Qy	61		GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC	120
Db	239		GCACCAACCATGGCCACGTTTGTGGAGCTCAGTACCAAAGCCAAGATGCCCATTTGTGGGC	298
Qy	121		CTGGGCACTTGGAAGTCTCCTCTCGGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT	180
Db	299		CTGGGCACTTGGAAGTCTCCTCTTGCAAAGTGAAAGAAGCAGTGAAGGTGGCCATTGAT	358
Qy	181		GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA	240
Db	359		GCAGGATATCGGCACATTGACTGTGCCTATGTCTATCAGAATGAACATGAAGTGGGGGAA	418
Qy	241		GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC	300
Db	419		GCCATCCAAGAGAAGATCCAAGAGAAGGCTGTGAAGCGGGAGGACCTGTTTCATCGTCAGC	478
Qy	301		AAGTTGTGGC-----	310
Db	479		AAGTTGTGGCCCACTTTCTTTGAGAGACCCCTTGTGAGGAAAGCCTTTGAGAAGACCCCTC	538
Qy	311	-----		310
Db	539		AAGGACCTGAAGCTGAGCTATCTGGACGTCTATCTTATTCACCTGGCCACAGGGATTCAAG	598
Qy	311	-----		310
Db	599		TCTGGGGATGACCTTTTCCCCAAAGATGATAAAGGTAATGCCATCGGTGGAAAAGCAACG	658
Qy	311	-----		310
Db	659		TTCTTGGATGCCTGGGAGGCCATGGAGGAGCTGGTGGATGAGGGGCTGGTGAAAGCCCTT	718
Qy	311	-----	CCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAA	353
Db	719		GGGGTCTCCAATTTTCAGCCACTTCCAGATCGAGAAGCTCTTGAACAAACCTGGACTGAAA	778
Qy	354		TATAAACCAGTGACTAACCAGGTTGAGTGTACCCATACCTCACGCAGGAGAAACTGATC	413
Db	779		TATAAACCAGTGACTAACCAGGTTGAGTGTACCCATACCTCACGCAGGAGAAACTGATC	838
Qy	414		CAGTACTGCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGAT	473
Db	839		CAGTACTGCCACTCCAAGGGCATCACCGTTACGGCCTACAGCCCCCTGGGCTCTCCGGAT	898
Qy	474		AGACCTTGGGCCAAGCCAGAAGACCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATT	533
Db	899		AGACCTTGGGCCAAGCCAGAAGACCTTCCCTGCTGGAGGATCCCAAGATTAAGGAGATT	958
Qy	534		GCTGCAAAGCACAAAAAAACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAAT	593
Db	959		GCTGCAAAGCACAAAAAAACCGCAGCCCAGGTTCTGATCCGTTTCCATATCCAGAGGAAT	1018
Qy	594		GTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTACAGGTCTTT	653
Db	1019		GTGATTGTCATCCCCAAGTCTGTGACACCAGCACGCATTGTTGAGAACATTACAGGTCTTT	1078
Qy	654		GACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGG	713
Db	1079		GACTTTAAATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGG	1138

us-10-653-681b-1.rni

Qy	714	GCCTGTAACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAAATATTGA	773
Db	1139	GCCTGTAACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCAATGCAGAAATATTGA	1198
Qy	774	GGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTAC	833
Db	1199	GGTTGAATCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTAC	1258
Qy	834	CTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAAGTGTGACTAC	893
Db	1259	CTCCACTCATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAAGTGTGACTAC	1318
Qy	894	ATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAG	953
Db	1319	ATAGACGAGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAG	1378
Qy	954	AATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATC	1013
Db	1379	AATATCACAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATC	1438
Qy	1014	AGAACAAATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATA	1073
Db	1439	AGAACAAATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATA	1498
Qy	1074	AAAAAAATAATAATCAT	1090
Db	1499	AAAAAAATAATAATCAT	1515

us-10-653-681b-1.rnpb

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 9, 2005, 06:00:04 ; Search time 1792 Seconds
(without alignments)
3942.932 Million cell updates/sec

Title: US-10-653-681B-1
Perfect score: 1090
Sequence: 1 caaaaacagcaacagaaagc.....ataaaaaaaataataatcat 1090

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7297361 seqs, 3241162794 residues

Total number of hits satisfying chosen parameters: 14594722

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

us-10-653-681b-1.rnpb

Result No.	Score	Query Match	Length	DB	ID	Description
1	1090	100.0	1090	21	US-10-653-681A-1	Sequence 1, Appli
2	832.2	76.3	1508	9	US-09-925-299-150	Sequence 150, App
3	832.2	76.3	1508	10	US-09-925-299-150	Sequence 150, App
4	796.6	73.1	1316	9	US-09-736-457-323	Sequence 323, App
5	796.6	73.1	1316	9	US-09-902-941-323	Sequence 323, App
6	796.6	73.1	1316	9	US-09-849-626-323	Sequence 323, App
7	796.6	73.1	1316	10	US-09-476-300-323	Sequence 323, App
8	796.6	73.1	1316	14	US-10-017-754-323	Sequence 323, App
9	796.6	73.1	1316	15	US-10-102-524-1731	Sequence 1731, Ap
10	796.6	73.1	1316	16	US-10-113-872-323	Sequence 323, App
11	796.6	73.1	1316	17	US-10-283-017-323	Sequence 323, App
12	718	65.9	1315	15	US-10-274-375-1	Sequence 1, Appli
13	707	64.9	1279	21	US-10-653-681A-3	Sequence 3, Appli
14	635.4	58.3	1816	14	US-10-198-846-14071	Sequence 14071, A
15	431	39.5	951	15	US-10-274-694-36	Sequence 36, Appl
16	431	39.5	951	20	US-10-332-448-36	Sequence 36, Appl
17	360.4	33.1	364	10	US-09-803-719-2393	Sequence 2393, Ap
18	353.4	32.4	356	10	US-09-803-719-2391	Sequence 2391, Ap
19	316	29.0	585	9	US-09-735-705-92	Sequence 92, Appl
20	316	29.0	585	9	US-09-850-716A-92	Sequence 92, Appl
21	316	29.0	585	9	US-09-897-778-92	Sequence 92, Appl
22	316	29.0	585	10	US-09-466-396A-92	Sequence 92, Appl
23	316	29.0	585	14	US-10-007-700-92	Sequence 92, Appl
24	316	29.0	585	15	US-10-117-982-92	Sequence 92, Appl
25	316	29.0	585	17	US-10-313-986-92	Sequence 92, Appl
26	316	29.0	585	20	US-10-775-972-92	Sequence 92, Appl
27	316	29.0	585	22	US-10-922-124-92	Sequence 92, Appl
C 28	316	29.0	858	9	US-09-735-705-91	Sequence 91, Appl
C 29	316	29.0	858	9	US-09-850-716A-91	Sequence 91, Appl
C 30	316	29.0	858	9	US-09-897-778-91	Sequence 91, Appl
C 31	316	29.0	858	10	US-09-466-396A-91	Sequence 91, Appl
C 32	316	29.0	858	14	US-10-007-700-91	Sequence 91, Appl
C 33	316	29.0	858	15	US-10-117-982-91	Sequence 91, Appl
C 34	316	29.0	858	17	US-10-313-986-91	Sequence 91, Appl
C 35	316	29.0	858	20	US-10-775-972-91	Sequence 91, Appl
C 36	316	29.0	858	22	US-10-922-124-91	Sequence 91, Appl
37	304.8	28.0	1331	20	US-10-737-450-97	Sequence 97, Appl
38	304.8	28.0	1335	18	US-10-641-643-1010	Sequence 1010, Ap
39	304.8	28.0	1367	9	US-09-864-864-332	Sequence 332, App
40	304.8	28.0	1367	15	US-10-171-581-1	Sequence 1, Appli
41	304.8	28.0	1367	17	US-10-172-118-641	Sequence 641, App
42	304.8	28.0	1367	18	US-10-342-887-641	Sequence 641, App
43	304.8	28.0	1367	19	US-10-717-597-289	Sequence 289, App
44	304.8	28.0	1368	19	US-10-416-330-21	Sequence 21, Appl
45	304.8	28.0	1416	21	US-10-887-553A-128	Sequence 128, App

ALIGNMENTS

RESULT 1

US-10-653-681A-1

; Sequence 1, Application US/10653681A

; Publication No. US20050048503A1

; GENERAL INFORMATION:

; APPLICANT: DAI, KEN-SHOW

; TITLE OF INVENTION: HUMAN ARL-RELATED GENE VARIANTS ASSOCIATED WITH CANCER

; FILE REFERENCE: U014798-3

; CURRENT APPLICATION NUMBER: US/10/653,681A

; CURRENT FILING DATE: 2003-09-02

```

; NUMBER OF SEQ NOS: 4
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
;   LENGTH: 1090
;   TYPE: DNA
;   ORGANISM: ARTIFICIAL SEQUENCE
;   FEATURE:
;   OTHER INFORMATION: VARIANT OF HUMAN ALDOSE REDUCTASE-LIKE GENE
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (70)..(333)
US-10-653-681A-1

```

Query Match 100.0%; Score 1090; DB 21; Length 1090;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Page 3

us-10-653-681b-1.rnpb

Db	661	AATTGAGTGATGAGGAGATGGCAACCATACTCAGCTTCAACAGAACTGGAGGGCCTGTA	720
Qy	721	ACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAA	780
Db	721	ACGTGTTGCAATCCTCTCATTTGGAAGACTATCCCTTCGATGCAGAATATTGAGGTTGAA	780
Qy	781	TCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACT	840
Db	781	TCTCCTGGTGAGATTATACAGGAGATTCTCTTTCTTCGCTGAAGTGTGACTACCTCCACT	840
Qy	841	CATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACG	900
Db	841	CATGTCCCATTTTAGCCAAGCTTATTTAAGATCACAGTGAACCTTAGTCCTGTTATAGACG	900
Qy	901	AGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCA	960
Db	901	AGAATCGAGGTGCTGTTTTAGACATTTATTTCTGTATGTTCAACTAGGATCAGAATATCA	960
Qy	961	CAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAA	1020
Db	961	CAGAAAAGCATGGCTTGAATAAGGAAATGACAATTTTTTCCACTTATCTGATCAGAACAA	1020
Qy	1021	ATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAA	1080
Db	1021	ATGTTTATTAAGCATCAGAACTCTGCCAACACTGAGGATGTAAAGATCAATAAAAAAAAA	1080
Qy	1081	TAATAATCAT	1090
Db	1081	TAATAATCAT	1090